



STIC Search Report

Biotech-Chem Library

File Copy
10/042059
updated

STIC Database Tracking Number: 10/042059

TO: David Lamberston
Location: rem/2b79/2c70
Art Unit: 1636
Friday, April 01, 2005

Case Serial Number: 10/042059

From: Barb O'Bryen
Location: Biotech-Chem Library
Remsen 1a69
Phone: 571-272-2518

BB
barbara.obryen@uspto.gov

Search Notes

STIC-Biotech/ChemLib

148695

From: Lambertson, David
Sent: Wednesday, March 23, 2005 1:32 PM
To: STIC-Biotech/ChemLib
Cc: Lambertson, David
Subject: Search Request

Search Request

Examiner's Name:	David Lambertson
Examiner #:	79514
Art Unit:	1636
Room #:	Remsen 02B79
Mailbox room#:	Remsen 02C70
Phone:	(571) 272-0771
Results Format:	paper

Serial # 10/042059

Please Search:

Nucleic Acid and Protein databases for:

SEQ ID NO: 1 and SEQ ID NO: 2

Including:

1. Default Search.
2. Interference Search.

Thanks,
Dave.

GenCore version 5.1.6
 Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model.

Run on: March 31, 2005, 04:36:43 ; Search time 99.5 Seconds

(without alignments) 1264.908 Million cell updates/sec

Title: US-10-042-059B-1

Perfect score: 1504

Sequence: atggactttatgaagccaga.....tctctcttggtgttacatg 843

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
 Ygapop 10.0 , Ygapext 0.5
 Fgapop 6.0 , Fgapext 7.0
 Delop 6.0 , Delext 7.0

Searched: 513545 seqs, 7464964 residues

Total number of hits satisfying chosen parameters: 1027090

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
 listing first 45 summaries

Command line parameters:

```
-MODEL=frame_n2p.model -DEV=xlh
-Q=100
-DB=CGN2.1/USP100.spc01/US1002059/runat_30032005_103354_7738/app_query.fasta_1.1031
-DBISSUED_PATENTS_AA=-OPMT=fastan -SUPPIX=xai -MINMATCH=0.1 -LOCPCII=0
-LOOPEXT=0 -UNITS=512 -START1=END1 -MATRIX=blosum62 -TRANS=human40.cdd
-LIST=45 -DOCNAME=200 -THR SCOREPCT -THR MAX=100 -THR MIN=0 -ALIGNS=15
-MODEL=LOCAL -OUTFILE=pto -NORM=ext -HEAVYSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10042055@CGN_1 -30 @runat_30032005_103354_7738 -NCPU=6 -ICPU=3
-NO_MMAP -LARGE_QUERY -NSEQ_SCORES=0 -WAIT -DBBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database : Issued_Patents_AA:*

```
1: /cgn2_6/podata/1/1/aa/5A-COMB.pep:*
2: /cgn2_6/podata/1/1/aa/5B-COMB.pep:*
3: /cgn2_6/podata/1/1/aa/6A-COMB.pep:*
4: /cgn2_6/podata/1/1/aa/6B-COMB.pep:*
6: /cgn2_6/podata/1/1/aa/backfiles1.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	921	61.2	315	4	US-09-248-796A-18213
2	508.5	33.8	257	4	US-09-454-219-14
3	481.5	32.0	331	4	US-09-454-219-12
4	464	30.9	334	4	US-09-454-219-21
5	423	28.1	261	4	US-09-454-219-16
6	399	26.5	265	4	US-09-454-219-22
7	313	20.8	207	4	US-09-454-219-18
8	221	14.7	154	4	US-09-454-219-8
9	197.5	13.1	132	4	US-09-454-219-6
10	132	8.8	93	4	US-09-454-219-2
11	127	8.4	248	4	US-09-252-911A-28060
12	62	4	US-09-454-219-4		

RESULT 1
 US-09-248-796A-18213
 ; Sequence 18213, Application US/09248796A
 ; Patent No. 6747337

GENERAL INFORMATION:

APPLICANT: Keith Weinstock et al

TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICINA

FILE REFERENCE: 107196.132

CURRENT APPLICATION NUMBER: US/09/248,796A

CURRENT FILING DATE: 1999-01-12

PRIOR APPLICATION NUMBER: US 60/074,725

PRIOR FILING DATE: 1998-02-13

PRIOR APPLICATION NUMBER: US 60/096,409

PRIOR FILING DATE: 1998-08-13

NUMBER OF SEQ ID NOS: 28208

SBO ID NO 18213

LENGTH: 315

TYPE: PRT

ORGANISM: Candida albicans

US-09-248-796A-18213

Alignment Scores:

Pred. No.: 4.76e-99

Score: 921.00

Percent Similarity: 82.63%

Best Local Similarity: 65.64%

Query Match: 61.24%

DB: 4

Length: 315

Matches: 170

Conservative: 44

Mismatches: 45

Indels: 0

Gaps: 0

US-10-042-059B-1 (1-843) x US-09-248-796A-18213 (1-315)

Qry 1 ATGGACTTTATGANGCCAGAACGCTGTGGACACATTAGAGATGCCTTGCAC

Db 48 MetAappheMetLysProgluThrValLeuAspLeuAlaAsnLeArginAlaLeuVal 67

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 31, 2005, 05:46:38 ; Search time 275 seconds
(Without alignments)
2032.966 Million cell updates/sec

Title: US-10-042-059B-1

Perfect score: 1504

Sequence: 1 atggactttatgaagccaga.....tccctcttggttgacttag 843

Scoring table: BLOSUM62

Xgapext 0.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delect 7.0

Searched: 1413372 seqs, 33152847 residues

Total number of hits satisfying chosen parameters: 2826744

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0*

Maximum Match 100*

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p,model -DEV=xlh
-Q/-cgn2_1/USFPO_spool/US1042059/runat_30032005_103356_7783/app_query.fasta_1.1031

-DB=Published Applications AA -QFM=rfabian -SUFFIX=rapb -MINMATCH=0.1

-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62

-TRANS=human40_d1 -LIS=45 -DOCALIGN=200 -TTR SCORE=pct -TTR MAX=100

-THR MIN=0 -ALIGN=1 -MODE=LOCAL -OUTFM=pfo -NORM=ext -HEAITSIZE=500 -MINLEN=0

-MAXLEN=200000000 -USERUS1042059 @CGN 1.1 103 @runat_30032005_103356_7783

-NCPU=6 -ICPU=3 -NO-NMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100

-LONGLOG -DEV TIMEOUT=120 -WARM TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5

-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELECT=7

Database : Published Applications AA:*

1: /cgn2_6/prodata/1/pubpa/US07_PUBCOMB.pep:*

2: /cgn2_6/prodata/1/pubpa/PCT_NW_PUB_PEP:*

3: /cgn2_6/prodata/1/pubpa/US06_NBW_PUB_PEP:*

4: /cgn2_6/prodata/1/pubpa/US06_PUBCOMB_PEP:*

5: /cgn2_6/prodata/1/pubpa/US07_NEW_PUB_PEP:*

6: /cgn2_6/prodata/1/pubpa/PCTUS_PUBCOMB_PEP:*

7: /cgn2_6/prodata/1/pubpa/US08_NEW_PUB_PEP:*

8: /cgn2_6/prodata/1/pubpa/US08_PUBCOMB_PEP:*

9: /cgn2_6/prodata/1/pubpa/US09_PUBCOMB_PEP:*

10: /cgn2_6/prodata/1/pubpa/US09_PUBCOMB_PEP:*

11: /cgn2_6/prodata/1/pubpa/US09_PUBCOMB_PEP:*

12: /cgn2_6/prodata/1/pubpa/US09_NEW_PUB_PEP:*

13: /cgn2_6/prodata/1/pubpa/US09_PUBCOMB_PEP:*

14: /cgn2_6/prodata/1/pubpa/US10_PUBCOMB_PEP:*

15: /cgn2_6/prodata/1/pubpa/US10_PUBCOMB_PEP:*

16: /cgn2_6/prodata/1/pubpa/US10_PUBCOMB_PEP:*

17: /cgn2_6/prodata/1/pubpa/US10_NEW_PUB_PEP:*

18: /cgn2_6/prodata/1/pubpa/US11_NEW_PUB_PEP:*

19: /cgn2_6/prodata/1/pubpa/US60_NEW_PUB_PEP:*

20: /cgn2_6/prodata/1/pubpa/US60_PUBCOMB_PEP:*

No.	Score	Match	Length	DB	ID	Description
1	1434	95.3	280	13	US-10-042-059A-2	Sequence 2, Appl
2	921	61.2	268	17	US-10-742-849-7032	Sequence 7, Appl
3	508.5	33.8	257	9	US-09-454-279-14	Sequence 14, Appl
4	508.5	33.8	257	15	US-10-624-061-14	Sequence 14, Appl
5	508.5	33.8	313	16	US-10-437-961-196439	Sequence 196339, Appl
6	497.5	33.1	320	15	US-10-425-114-6585	Sequence 6886, Appl
7	495.5	32.9	315	15	US-10-424-599-266662	Sequence 266662, Appl
8	486	32.3	350	16	US-10-767-701-45917	Sequence 45917, Appl
9	481.5	32.0	331	9	US-09-454-279-12	Sequence 12, Appl
10	481.5	32.0	331	15	US-10-624-061-12	Sequence 12, Appl
11	464	30.9	334	9	US-09-454-279-21	Sequence 21, Appl
12	454	30.9	334	15	US-10-624-061-21	Sequence 21, Appl
13	423	28.1	261	9	US-09-454-279-15	Sequence 16, Appl
14	423	28.1	261	15	US-10-424-061-16	Sequence 16, Appl
15	423	28.1	297	15	US-10-425-114-48703	Sequence 48703, Appl
16	419	27.9	261	15	US-10-424-599-235878	Sequence 235878, Appl
17	408.5	27.2	255	16	US-09-454-279-22	Sequence 151855, Appl
18	399	25.5	265	9	US-09-454-279-22	Sequence 22, Appl
19	399	25.5	265	15	US-10-624-061-22	Sequence 22, Appl
20	374.5	24.9	313	15	US-10-425-114-65608	Sequence 65608, Appl
21	373	24.8	264	15	US-10-425-114-65528	Sequence 65528, Appl
22	356	23.7	306	16	US-10-437-963-186047	Sequence 186047, Appl
23	348	23.1	316	16	US-10-425-114-386040	Sequence 186040, Appl
24	344.5	22.9	186	15	US-10-425-114-52777	Sequence 52677, Appl
25	328	21.8	199	15	US-10-424-114-43606	Sequence 43606, Appl
26	328	21.8	207	9	US-09-454-279-18	Sequence 18, Appl
27	313	20.8	207	15	US-10-624-061-18	Sequence 103185, Appl
28	297.5	19.8	489	16	US-10-437-963-103185	Sequence 103185, Appl
29	282	18.8	210	16	US-10-424-061-32934	Sequence 186047, Appl
30	221	14.7	154	9	US-09-454-279-8	Sequence 8, Appl
31	221	14.7	154	15	US-10-624-061-8	Sequence 6, Appl
32	197.5	13.1	132	15	US-10-624-061-6	Sequence 6, Appl
33	197.5	13.1	132	15	US-10-624-061-6	Sequence 6, Appl
34	178	11.8	109	15	US-10-424-599-224668	Sequence 224668, Appl
35	17.5	10.5	97	15	US-10-424-599-224676	Sequence 38176, Appl
36	151.5	10.1	19695	15	US-10-0-846A-3	Sequence 3, Appl
37	132	8.8	93	9	US-09-454-279-2	Sequence 2, Appl
38	132	8.8	93	15	US-0-0-621-1	Sequence 2, Appl
39	130.5	8.8	278	15	US-10-425-114-48480	Sequence 48480, Appl
40	130.5	8.9	279	15	US-10-425-114-43408	Sequence 43408, Appl
41	124.5	8.3	797	14	US-09-454-279-10907	Sequence 10907, Appl
42	122	8.1	62	9	US-09-454-279-4	Sequence 4, Appl
43	122	8.1	62	15	US-10-624-061-4	Sequence 4, Appl
44	119	7.9	182	16	US-10-437-963-116837	Sequence 116837, Appl
45	117.5	7.8	459	16	US-10-437-963-153917	Sequence 153917, Appl

ALIGNMENTS

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result

Query

RESULT 1
US-10-042-059A-2
Publication No. US20020197704A1
GENERAL INFORMATION:
APPLICANT: Rhein Biotech Gesellschaft fur neue biotechnologische Prozesse und
TITLE OF INVENTION: Nucleic Acid Molecule Containing a Nucleic Acid Coding for a Pol
TITLE OF REFERENCE: PCT/IB98/05803
CURRENT APPLICATION NUMBER: US10/042,059A
CURRENT FILING DATE: 2001-10-25
PRIOR APPLICATION NUMBER: DE 199 19 124 7
PRIOR FILING DATE: 1999-04-27
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 280
TYPE: PRT
ORGANISM: Hansenula polymorpha
US-10-042-059A-2

Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on:

March 31, 2005, 04:14:20 ; (without alignments)

3686.851 Million cell updates/sec

Title: US-10-042-059B-1

Perfect score: 1504

Sequence: 1 atggactttatgaagccaga.....tctccatctgggttgacttag 843

Scoring table: blosum62

Xgapop 10.0 , Xgapext 0.5

Xgapop 10.0 , Xgapext 0.5

Xgapop 6.0 , Xgapext 7.0

Xgapop 6.0 , Xgapext 7.0

Deletx 6.0 , Deletx 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+npb:modelf -DEV=xlh
-Q=/sgm2/1/USPRO_SPOOL/US10042059/runat_30032005_103353_7727/app_query.fasta_1.1031
-DB=ppr_79 -QFM=Fastan
-UNITS=5bit -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40_cdi -LIST=45 .
-DOCALIGN=200 -THR_SCORE=perc -THR_MAX=100 -MINLEN=0 -ALIGNN=15 -MODE=LOCAL
-OUTFILE=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10042059 @CGN 1.1.25 @runat_30032005_103353_7727_NCPD=6 -ICPU=3
-NO_MMDF -LARGEQTER=1 -NEG_SCORES=100 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=1200 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database : PIR_79:*

```
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

RESULT 1

Ad5921

chorimate mutase (EC 5.4.99.5) - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein YPP499.15C; protein YPR060c

C:Species: Saccharomyces cerevisiae

C:Accession: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004

C:Accession: Ad5921; S54081

J: Bacteriol. 171: 1245-1253, 1989

A:Title: A single point mutation results in a constitutively activated and feedback-regulated gene

A:Reference number: Ad5921; MUID:89155418; PMID:2646272

A:Accession: Ad5921

A:Molecule type: DNA

A:Residues: 1-256 <SGH>

A:Cross-references: UNIPROT:P32178; GB:M24517; NID:9295576; PIDN:AAB59309.1; PID:929557

R:Badcock, K., Churcher, C.M. Library, May 1995

A:Reference number: S54059

A:Accession: S54081

A:Molecule type: DNA

A:Residues: 1-256 <BRD>

A:Cross-references: SGD:Z49219; NID:9805025; PIDN:CAA89177.1; PID:9805040; MIPS:YPRO6C

A:Experimental source: Strain AB972

C:Genetics:

A:Gene: SGD:AR07

A:Cross-references: SGD:S0006264; MIPS:YPR060c

A:Map position: 16R

C:Function:

A:Description: intramolecular transferase; isomerase

A:Pathway: aromatic amino acid biosynthesis

C:Superfamily: chorimate mutase of the AroQ class, eukaryotic type

C:Keywords: aromatic amino acid biosynthesis; intramolecular transferase; isomerase

Alignment Scores:

Pred. No.: 6.12e-55 Length: 256

Score: 776.50 Matches: 152

Percent Similarity: 74.13% Conservative: 40

Best Local Similarity: 58.69% Mismatches: 62

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

On nucleic - protein search, using frame_plus_n2p model

Run on:

March 31, 2005, 03:04:01 ; Search time 247.5 Seconds

{without alignment} 3488.345 Million cell updates/sec

Title: US-10-042-059B-1

Perfect score: 1504

Sequence: 1 atggactttatgaagccaga.....tccctctgggtttag 843

Scoring table: BLOSUM2

Xgapop 1.0.0 Xgapext 0.5
Ygapop 1.0.0 Ygapext 0.5
Fgapop 6.0 Fgapext 7.0
Delop 6.0 Delect 7.0

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 3224756

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```
-MODEU=frame+n2p;model -DEV=xlh
-Q=/CGN2_1/USPTO/spool/US10042059/runat_30032005_103352_7710/app_query.fasta_1.1031
-DB=UniProt_03 -FORMAT=fasta -SEARCHTYPE=BLAST -MINMATCH=0.1 -LOOPCUT=0 -LOOPCUTN=45
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=95 -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFILE=pt0 -NORM=ext -HEAPSIZE=500M -MINLEN=0 -MAXLEN=2000000000
-USER-US10042059 @CGN_1_1_140@runat_30032005_103352_7710 -NCPU=6 -ICPU=3
-NO_MMAPP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPEXT=10 -XGAPEXT=0.5 -FQAPOP=6
-FQAPEXT=7 -YGAPOD=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

RESULT:1

Q9P4DB

ID

OP4DB

PRELIMINARY

PRT

280 AA.

OP4DB8;

ID

01-OCT-2000

(Tremblrel.

15, Created)

DT

01-OCT-2004

(Tremblrel.

15, Last sequence update)

DT

01-MAR-2004

(Tremblrel.

15, Last annotation update)

DT

Chorismate mutase (EC 5.4.99.5).

GN

Name=HAK07;

Pichia angusta (Yeast) (Hansenula polymorpha).

OS

Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC

Saccharomycetidae; Saccharomycetaceae; Pichia.

OX

NCBI - TaxID=4905;

[1]

RN

SEQUENCE FROM N.A.

RC

STRAIN=RB11;

RX

MEDLINE=20033451; PubMed=10894726;

DOI=10.1128/JB.082.15.4118-4197.2000;

Krappmann S., Pries R., Gellissen G., Hiller M., Braus G.H.: "HAK07 encodes chorismate mutase of the methylotrophic yeast Hansenula polymorpha and is derepressed upon methanol utilization."; J. Bacteriol. 182:4188-4197 (2000).

RL

EMBL; AF204738; AAF87954.1;

DR

HSSP; P32178; 2CSM.

DR

GO; GO:000106; F:chorismate mutase activity; IEA.

DR

GO; GO:0015853; F:isomerase activity; IEA.

DR

GO; GO:0005073; F:chorismate mutase activity; IEA.

DR

InterPro; IPR002701; Chorismate mutase activity; IEA.

DR

InterPro; IPR008233; Chorismate mutase activity; IEA.

DR

Pfam; PF01817; CM_2; 1.

DR

PIR; PIR017118; Chorismate mutase activity; IEA.

DR

TIGRFams; TIGR01802; CM_2; 1.

ALIGNMENTS

Q9x660 arabidopsis
Q6h819 oryza sativ
Q957h4 arabidopsis
Q659e6 oryza sativ
Q65311 nicotiana
Q55861 prunus avium
Q2410 petroelium
Q2249 petroelium
Q9vr49 drosophila
Q51102 borrelia bu
Q651h3 bacillus li
Q8dc7 xenopus lae
Q8tib8 mus musculus
Q8djr3 arabidopsis
Q9ci05 schizosaccharomyces
P4014 drosophila
Q19791 caenorhabditis
Q1zpe6 mycoplasma
Q833p7 mus musculus
Q9dss1 rat cytomegalovirus
Q9csl1 mus musculus
Q8v4z4 monkeypox virus
Q9pdil mus musculus
Q9zpe9 mus musculus
Q9q4z4 mus musculus
Q8z03 lactobacillus
Q9697 homo sapiens
Q8iwq7 homo sapiens
Q9nwz1 homo sapiens

Q9x660 arabidopsis
Q6h819 oryza sativ
Q957h4 arabidopsis
Q659e6 oryza sativ
Q65311 nicotiana
Q55861 prunus avium
Q2410 petroelium
Q2249 petroelium
Q9vr49 drosophila
Q51102 borrelia bu
Q651h3 bacillus li
Q8dc7 xenopus lae
Q8tib8 mus musculus
Q8djr3 arabidopsis
Q9ci05 schizosaccharomyces
P4014 drosophila
Q19791 caenorhabditis
Q1zpe6 mycoplasma
Q833p7 mus musculus
Q9dss1 rat cytomegalovirus
Q9csl1 mus musculus
Q8v4z4 monkeypox virus
Q9pdil mus musculus
Q9zpe9 mus musculus
Q9q4z4 mus musculus
Q8z03 lactobacillus
Q9697 homo sapiens
Q8iwq7 homo sapiens
Q9nwz1 homo sapiens

Percent Similarity: 100.00%
Best Local Similarity: 100.00%
MisMatches: 0

Copyright (c) 1993 - 2005 Compugen Ltd.

Om nucleic - nucleic search, using sw model

Run on: March 30, 2005, 21:05:55 ; Search time 4115 Seconds
(without alignments)
9926.548 Million cell updates/sec

Title: US-10-042-059B-1

Perfect score: 843

Sequence: 1 atggactttatgaagcaga.....tctccctctgggttgactag 843

Scoring table: IDENTITY_NUC

Gapext 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 10%

Listing first 45 summaries

Database :

Alignments

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	843	100.0	843	BD266599
2	843	100.0	843	BD266599
3	843	100.0	1652	BD266599
4	843	100.0	1655	BD266600
5	843	100.0	1655	BD266600
6	325.2	38.6	110000	BD266760
7	281.8	33.4	110000	CR382138_20
8	270	32.0	948	AR448979
9	256.4	30.4	110000	AB016817_01
10	217.4	25.8	2059	YSCAR07A
11	217.4	25.8	43776	SC9499X
12	217.4	25.8	165311_20	SCCHXVI
13	217.2	25.8	771	AY633179
14	210.4	25.0	110000	CR382122_07
15	189	22.4	110000	CR380957_11
16	156.4	18.6	869	BD178335
17	156.4	18.6	869	AB116236
18	156.4	16.8	1348	AK105512
19	131.4	15.6	1231	AR404687

RESULT 1

BD266599

LOCUS BD266599

DEFINITION Nucleic acid molecule, containing a nucleic acid which codes for a polypeptide with chorismate mutase activity.

ACCESSION BD266599

VERSION BD266599-1 GR:33076367

KEYWORDS JP 2002542788-A/1.

SOURCE Pichia angusta

ORGANISM Pichia angusta

Bukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Pichia.

REFERENCE 1 (bases 1 to 843)

AUTHORS Gellissen, G., Braus, G., Pries, R., Krappmann, S. and Strasser, A.W.

TITLE Nucleic acid molecule, containing a nucleic acid which codes for a polypeptide with chorismate mutase activity

JOURNAL Patent: JP 2002542788-A 17-DEC-2002; RHEIN BIOTECH GESELLSCHAFT FÜR NEUE BIOTECHNOLOGISCHE PROZESSE UND PRODUKTE MBH

COMMENT OS Hansenula polymorpha (Yeast)

PN JP 2002542788-A/1

PD 17-DEC-2002

PP 27-APR-2000 JP 2000614405

PR 27-APR-1999 DB 199 19 124-7

PI GERD GELLISEN, GERHARD BRAUS, RALPH PRIES, SVEN KRAPPmann, PI ALEXANDER W STRASSER

PC C12N15/09, C07K16/40, C12N1/15, C12N1/19, C12N1/21, C12N5/110

;C12N9/90, C12N15/00,

PC C12N5/00

CC Nucleic acid molecule, containing a nucleic acid which codes

CC for a

CC polypeptide with chorismate mutase activity

PH Location/Qualifiers

FT Key

FT source 1..843

FEATURES /organism="Hansenula polymorpha (Yeast)".

source 1..843

/organism="Pichia angusta"

/mol_type="genomic DNA"

/db_xref="taxon:4905"

ORIGIN AR404687 Sequence

Copyright (c) 1993 - 2005 Compugen Ltd.

GanCore version 5.1.6

Om nucleic - nucleic search, using SW model

Run on: March 30, 2005, 18:17:40 ; Search time 588 Seconds
(without alignments)

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: US-10-042-059B-1
843 1 atggactttatggacccaga.....tctccatctggctgtactag 843
8486.976 Million cell updates/sec

Perfect score: 843

Sequence: IDENTITY_NUC

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq 16Dec04:
1: geneseqn1980s:
2: geneseqn1990s:
3: geneseqn2000s:
4: geneseqn2001as:
5: geneseqn2001bs:
6: geneseqn2002as:
7: geneseqn2002bs:
8: geneseqn2003as:
9: geneseqn2003bs:
10: geneseqn2003c:
11: geneseqn2003ds:
12: geneseqn2004as:
13: geneseqn2004bs:
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	843	100.0	843 3 AAC81949	AAC81949 H. polymo Aac81950 H. polymo
2	843	100.0	1655 3 AAC81950	Adp9547 C. albica
3	270	32.0	807 12 ADP98547	Adrp8510 Aspergillus
4	182.8	21.7	801 13 ADP8510	Aal50202 M. sterilis
5	156.4	18.6	869 8 AAL50202	Adr8923 Aspergillus
6	141.2	16.7	989 13 ADP84933	Abt17891 Aspergillus
7	141.2	16.7	2920 8 ABT17891	Abt19705 Aspergillus
8	141.2	16.7	6989 13 ADP84336	Adr8336 Aspergillus
9	131.4	15.6	1231 10 ABX93057	Abx93057 Rice chor
10	131.4	15.6	1231 12 ADJ5263	Adj5263 Rice chor
11	131.4	15.6	1231 12 ADJ5263	Adj5263 Rice chor
12	116.8	13.9	658 3 ADP14469	Aaf1469 Aspergillus
13	111.2	13.2	1223 10 ABX93056	Abx93056 Corn chor
14	111.2	13.2	1223 12 ADJ5261	Adj5261 Corn chor
15	104.4	12.4	1217 6 ABX91395	Aba91395 Arabidops
16	104.4	12.4	1217 6 ABX91401	Aba91401 Arabidops
17	99.8	11.8	1020 10 ABX91058	Abx91058 Soybean C
18	99.8	11.8	1020 12 ADJ55265	Adj55265 Soybean C
19	97.4	11.6	798 6 ABZ13167	Abz13167 Arabidops
20	97.4	11.6	993 6 ABX99614	Aba99614 Construct

ALIGNMENTS

Result No.	Score	Query Match Length	DB ID	Description
1	97.4	11.6	1006 6 ABA99613	Ab99613 A. thalia
2	97.4	11.6	1006 6 ABA91400	Ab91400 Arabidops
3	97.4	11.6	1006 6 ABA91394	Ab91394 Arabidops
4	93.6	11.1	1005 6 ABZ13232	Abz13232 Arabidops
5	93.6	11.1	1207 6 ABA91393	Ab91393 Arabidops
6	93.6	11.1	1207 6 ABA91399	Ab91399 Arabidops
7	93.6	11.1	1274 3 AAC3324	Aac3324 Arabidops
8	93.6	11.1	580 13 ACN51510	Acn51510 Cotton gy
9	69.8	8.3	622 13 ACN50132	Acn50132 Cotton no
10	64.2	7.6	600 13 ADR64815	Adr64815 Cotton CD
11	61.8	7.3	258 6 ABL71526	Ab71526 Corn tass
12	61	7.2	780 12 ADJ5267	Adj5267 Wheat cho
13	61	7.2	780 12 ADJ5267	Adj5267 Wheat cho
14	59.8	7.1	579 10 ABX91054	Abx91054 Wheat cho
15	59.8	7.1	579 12 ADJ5257	Adj5257 Wheat cho
16	56.2	6.7	584 13 ACN51205	Acn51205 Cotton no
17	54.8	6.5	542 13 ACN57601	Acn57601 Cotton gy
18	53.4	6.3	525 12 ADJ5255	Adj5255 Soybean c
19	53.4	6.3	525 12 ADJ5255	Adj5255 Soybean c
20	52.6	6.2	524 13 ADR64235	Adr64235 Cotton cd
21	49.4	5.9	2000 8 ADA71938	Ada71938 Rice gene
22	48.8	5.8	864 12 ADJ44958	Adj44958 plant cdn
23	47.6	5.6	1379 6 ABA91794	Ab91794 Yeast mit
24	47	5.6	5858 8 AAD55726	Ad55726 Nephila m
25	45	5.6	180385 10 ADL13931	Adl13931 Osteoarth

This invention describes novel nucleic acids (I) that encode a polypeptide (II) with chorismate mutase (CM) activity (or its complementarity strand). CM catalyzes conversion of chorismate to prephenate, an essential precursor for Phe and Tyr. (I) is a selection marker for construction of corresponding auxotrophic yeast (requiring Phe and Tyr) that are used for recombinant production of proteins. (I) allows selection of transformed yeast on simple media. (updated on 15-SEP-2003)

GenCore version 5.1.6
 Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using BW model

Run on: March 31, 2005, 00:11:41 ; Search time 195 Seconds
 (without alignments)
 7073.750 Million cell updates/sec

Title: US-10-042-059B-1
 Perfect score: 843
 Sequence: 1 atggactttatgaagcaga.....tctctctgggttgtaactag 843
 Scoring table: IDENTITY_NUC
 Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : ISSUED_PATENTS_NA:*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq: *
 6: /cgn2_6/ptodata/1/ina/backfileseq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

RESULT 1
 ; Sequence 4110, Application US/09248796A
 ; Patent No. 6747137
 ; GENERAL INFORMATION:
 ; APPLICANT: Keith Weinstock et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
 ; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.132
 ; CURRENT APPLICATION NUMBER: US/09/248,796A
 ; CURRENT FILING DATE: 1999-02-12
 ; PRIORITY APPLICATION NUMBER: US 60/074,725
 ; PRIORITY FILING DATE: 1998-02-13
 ; NUMBER OF SEQ ID NOS: 28208
 ; SEQ ID NO 4110
 ; LENGTH: 948
 ; TYPE: DNA
 ; ORGANISM: Candida albicans

US-09-248-796A-4110

Query Match 32.0%; Score 270; DB 4; Length 948;
 Best Local Similarity 59.1%; Pred. No. 1.4e-72;
 Matches 462; Conservative 0; Mismatches 320; Index 0; Gaps 0;

QY 1 ATGGACTTTATGAAAGCAGAACAGTCCTGCAACTTGTGAGATGCTTGTC 60
 Db 142 ATGGATTTTATGAAAGCAGAACAGTCCTGCAACTTGTGAGATGCTTGTA 201

QY 61 CGGATGGAGGATAGCATTTCACTTATGAGGGTCTTATCGTGGCC 120
 Db 202 AGGATGGAGGATACTATGTGTTGATTAAATCGAAGAGATCTCAATTTCAGTCCCCA 261

QY 121 TGGTATACAAAGTCACCCAGTCCTATTCCAACTTCAGGGCTGGTCTTGAACTG 180
 Db 262 TCAAGTTATGAAAGTAAATATATTCGCCATTGTGGACTTTTGGATGG 321

QY 181 CTTGGGGGAGCACCGGGATCCATTGAGGATACCAAGGCCAGGAG 240
 Db 322 GCTTGTGTTACATTGGAGGTGCTCATCTCAATGAGCTTATGAGGCCAGGAA 381

QY 241 GCGCTTCTTCCCAAGCTCTGAAAGAAGTTCCTGCCAAGTCAATTACCCATG 300
 Db 382 ACTCCATTTCAGTCATTGAAACTCCATTTCAGGCCATTATCTCAA 441

QY 301 GTCCTACCTCTACGGGATCAACCTCACAAAGAGATCTCAAGTCACCG 360

Db 442 ATATGGCAAATATTCGAGAAATTGAAATTGTTCTGAAATATGAGTTTGT 501

Sequence 15259, A
 Sequence 3, Appli
 Sequence 5, Appli
 Sequence 6, Appli
 Sequence 3, Appli
 Sequence 2, Appli
 Sequence 1, Appli
 Sequence 13635, A
 Sequence 13809, A
 Sequence 12701, A
 Sequence 13259, A
 Sequence 6, Appli
 Sequence 89, Appli
 Sequence 8, Appli
 Sequence 8, Appli
 Sequence 5061, Appli
 Sequence 20343, A
 Sequence 6442, Appli

Sequence 15259, A
 Sequence 3, Appli
 Sequence 5, Appli
 Sequence 6, Appli
 Sequence 3, Appli
 Sequence 2, Appli
 Sequence 1, Appli
 Sequence 13635, A
 Sequence 13809, A
 Sequence 12701, A
 Sequence 13259, A
 Sequence 6, Appli
 Sequence 89, Appli
 Sequence 8, Appli
 Sequence 8, Appli
 Sequence 5061, Appli
 Sequence 20343, A
 Sequence 6442, Appli

Copyright (c) 1993 - 2005 Compugen Ltd. GenCore version 5.1.6

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

ALIGNMENT

Sequence 7285, Ap
Sequence 102347, Ap
Sequence 4455, Ap
Sequence 11, Appl
Sequence 11, Appl
Sequence 14353, A
Sequence 102351,
Sequence 102346,
Sequence 4, Appl
Sequence 10, Appl
Sequence 102349,
Sequence 15, Appl
Sequence 15, Appl
Sequence 76339, Ap
Sequence 972, App
Sequence 972, App
Sequence 3, Appl
Sequence 9, Appl
Sequence 4454, Ap
Sequence 737, App
Sequence 737, App
Sequence 2, Appl
Sequence 8, Appl
Sequence 621, App
Sequence 93036, A
Sequence 3898, Ap
Sequence 151470, A
Sequence 27576, A
Sequence 102342,
Sequence 49372, A
Sequence 83557, A
Sequence 1332, Ap
Sequence 115559, A
Sequence 12291, A

Copyright (c) 1993 - 2005 Compugen Ltd.
GenCore version 5.1.6

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2005, 08:01:45 ; (without alignments)

Scoring table: BLOSUM62 (230.601 Million cell updates/sec)

Title: US-10-042-059B-2

Perfect score: 1434

Sequence: 1 MDFMKRPEVLDLGNIRDALV.....DDATQKSGGYVDRPLSGLV 280

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_16Dec04:*

- 1: geneseqp1980s:*
- 2: geneseqp2000s:*
- 3: geneseqp2001s:*
- 4: geneseqp2002s:*
- 5: geneseqp2003as:*
- 6: geneseqp2003bs:*
- 7: geneseqp2004as:*
- 8: geneseqp2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

ALIGNMENTS

26	221	15.4	154	8	ADJ55258	Adj55258 Wheat cho
27	197.5	13.8	132	6	ABU08089	Abu08089 Protein e
28	197.5	13.8	132	8	ADJ55256	Adj55256 Soybean c
29	132	9.2	93	8	ADJ55252	Adj55252 Corn chor
30	132	9.2	93	8	ADJ55252	Adj55252 Corn chor
31	122	8.5	62	6	ABU08088	Abu08088 Protein e
32	122	8.5	62	8	ADJ55254	Adj55254 Rice chor
33	114.5	8.0	446	2	ABG15384	Abg15384 Novel hum
34	114.5	8.0	469	2	ABG15384	Abg15384 Novel hum
35	10.5	7.4	1638	8	ADK71824	Adk71824 Human kin
36	103.5	7.2	1732	8	ADP95103	Adp95103 Rat. Berlin
37	103.5	7.2	1732	8	ADP89932	Adp89932 Rat. Berlin
38	103	7.2	724	4	ABG15384	Abg15384 Novel hum
39	103	7.2	724	4	ABG15384	Abg15384 Novel hum
40	10.2	7.1	552	6	ADR43675	Adr43675 Sarcina v
41	10.2	7.0	1664	7	ADP47740	Adp47740 Human Nov
42	10.5	7.0	1664	8	ADJ79010	Adj79010 Human Nov
43	10.5	7.0	1719	5	ADR21707	Aad21707 Human PK1
44	10.5	7.0	1732	8	ADP89931	Adp89931 Human ser
45	100.5	7.0	1732	8	ADS93536	Adp93536 Human MRC

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_16Dec04:*

- 1: geneseqp1980s:*
- 2: geneseqp2000s:*
- 3: geneseqp2001s:*
- 4: geneseqp2002s:*
- 5: geneseqp2003as:*
- 6: geneseqp2003bs:*
- 7: geneseqp2004as:*
- 8: geneseqp2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

ALIGNMENTS

RESULT 1

AB11451 ID AB11451 standard; protein; 280 AA.

AC AB11451; DT 12-SEP-2003 (revised) 01-MAR-2001 (first entry)

DR H. polymorpha chorismate mutase protein.

KW Chorismate mutase; prephenate; selection marker; auxotrophic yeast.

OS Pichia angusta.

XX PN WO20065071-A1.

XX PD 02-NOV-2000.

XX PR 27-APR-2000; 99DE-01019124.

XX PR (RHEIN-) RHEIN BIOTECH GES NEUE BIOTECHNOLOGISCHE.

XX PT Gellissen G, Braus G, Pries R, Krappmann S, Strasser AW;

XX DR WPI; 2000-687355/67.

DR N-PSDR; ADR81949.

This invention describes novel nucleic acids (I) that encode a polypeptide (II) with chorismate mutase (CM) activity (or its complementary strand). CM catalyzes conversion of chorismate to prephenate, an essential precursor for Phe and Tyr. (I) is a selection marker for construction of corresponding auxotrophic yeast (requiring Phe and Tyr) that are used for recombinant production of proteins. (I) allows selection of transformed yeast on simple media. (Updated on 12-SEP-2003 to standardise OS field)

XX SQ sequence 280 AA;

PS Claim 1h; Page 57-58; 63PP; German.

CC This invention describes novel nucleic acids (I) that encode a polypeptide (II) with chorismate mutase (CM) activity (or its complementary strand). CM catalyzes conversion of chorismate to prephenate, an essential precursor for Phe and Tyr. (I) is a selection marker for construction of corresponding auxotrophic yeast (requiring Phe and Tyr) that are used for recombinant production of proteins. (I) allows selection of transformed yeast on simple media. (Updated on 12-SEP-2003 to standardise OS field)

CC sequence 280 AA;

CC Query Match 100.0%; Score 1434; DB 3; Length 280;

25 Query Match 100.0%; Score 1434; DB 3; Length 280;

GenCore version 5.1.6 Copyright (c) 1993 - 2005 Compugen Ltd.									
OM protein - protein search, using SW model									
Run on: March 31, 2005, 10:14:39 ; Search time 29 Seconds (without alignments)									
Post-processing: Minimum Match 0%									
Maximum DB seq length: 0 Maximum DB seq length: 2000000000									
Listings first 45 summaries									
Database : Issued Patents AA:*									
1: /cgn2_6/ptodata/1/iaa/5A-COMB.pep:*									
2: /cgn2_6/ptodata/1/iaa/5B-COMB.pep:*									
3: /cgn2_6/ptodata/1/iaa/6A-COMB.pep:*									
4: /cgn2_6/ptodata/1/iaa/6B-COMB.pep:*									
5: /cgn2_6/ptodata/1/iaa/PCUTUS-COMB.pep:*									
6: /cgn2_6/ptodata/1/iaa/backfilesl1.pep:*									
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.									
SUMMARIES									
Result No. Score Query Length DB ID Description									
1	921	64.2	315	4	US-09-248-796A-18213	Sequence 18213, A	Sequence 18213, A	RESULT 1	US-09-248-796A-18213
2	508.5	35.5	45	4	US-09-248-796A-18213	Sequence 18213, A	Sequence 18213, A	Sequence 18213, A	Sequence 18213, A
3	481.5	33.6	331	4	US-09-454-279-12	Sequence 12, Appl	Sequence 12, Appl	Sequence 12, Appl	Sequence 12, Appl
4	464	32.4	334	4	US-09-454-279-21	Sequence 21, Appl	Sequence 21, Appl	Sequence 21, Appl	Sequence 21, Appl
5	423	29.5	261	4	US-09-454-279-16	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl
6	399	27.8	265	4	US-09-454-279-22	Sequence 22, Appl	Sequence 22, Appl	Sequence 22, Appl	Sequence 22, Appl
7	313	21.8	207	4	US-09-454-279-18	Sequence 18, Appl	Sequence 18, Appl	Sequence 18, Appl	Sequence 18, Appl
8	221	15.4	157	4	US-09-454-279-8	Sequence 8, Appl	Sequence 8, Appl	Sequence 8, Appl	Sequence 8, Appl
9	197.5	13.8	132	4	US-09-454-279-6	Sequence 6, Appl	Sequence 6, Appl	Sequence 6, Appl	Sequence 6, Appl
10	132	9.3	93	4	US-09-454-279-2	Sequence 2, Appl	Sequence 2, Appl	Sequence 2, Appl	Sequence 2, Appl
11	122	8.5	62	4	US-09-454-279-4	Sequence 4, Appl	Sequence 4, Appl	Sequence 4, Appl	Sequence 4, Appl
12	94	6.6	391	4	US-09-270-677-22309	Sequence 42309, A	Sequence 42309, A	Sequence 42309, A	Sequence 42309, A
13	94	6.6	502	4	US-09-635-872A-13	Sequence 13, Appl	Sequence 13, Appl	Sequence 13, Appl	Sequence 13, Appl
14	94	6.6	502	4	US-09-635-077A-13	Sequence 13, Appl	Sequence 13, Appl	Sequence 13, Appl	Sequence 13, Appl
15	94	6.6	502	4	US-09-636-060C-13	Sequence 13, Appl	Sequence 13, Appl	Sequence 13, Appl	Sequence 13, Appl
16	94	6.6	502	4	US-09-986-552-13	Sequence 13, Appl	Sequence 13, Appl	Sequence 13, Appl	Sequence 13, Appl
17	94	6.6	502	4	US-09-636-596C-13	Sequence 13, Appl	Sequence 13, Appl	Sequence 13, Appl	Sequence 13, Appl
18	94	6.6	502	4	US-10-023-894-16	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl
19	94	6.6	502	4	US-10-306-686-13	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl
20	92	6.4	717	4	US-09-583-110-4629	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl
21	92	6.4	721	4	US-09-107-433-2019	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl
22	90	6.3	292	4	US-09-107-532A-3741	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl
23	89.5	6.2	640	4	US-09-248-796A-16129	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl
24	89	6.2	335	4	US-09-934-901-6	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl
25	89	6.2	335	4	US-09-934-868-16	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl
26	89	6.2	335	4	US-10-321-2810-5	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl
27	89	6.2	335	4	US-10-320-8874-6	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl
Searched: 513545 seqs, 74659064 residues									
Total number of hits satisfying chosen parameters: 513545									
Minimum DB seq length: 0									
Maximum DB seq length: 2000000000									
Post-processing: Minimum Match 0%									
Listings first 45 summaries									
Database : Issued Patents AA:*									
1: /cgn2_6/ptodata/1/iaa/5A-COMB.pep:*									
2: /cgn2_6/ptodata/1/iaa/5B-COMB.pep:*									
3: /cgn2_6/ptodata/1/iaa/6A-COMB.pep:*									
4: /cgn2_6/ptodata/1/iaa/6B-COMB.pep:*									
5: /cgn2_6/ptodata/1/iaa/PCUTUS-COMB.pep:*									
6: /cgn2_6/ptodata/1/iaa/backfilesl1.pep:*									
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.									
SUMMARIES									
Result No. Score Query Length DB ID Description									
1	921	64.2	315	4	US-09-248-796A-18213	Sequence 18213, A	Sequence 18213, A	RESULT 1	US-09-248-796A-18213
2	508.5	35.5	45	4	US-09-248-796A-18213	Sequence 18213, A	Sequence 18213, A	Sequence 18213, A	Sequence 18213, A
3	481.5	33.6	331	4	US-09-454-279-12	Sequence 12, Appl	Sequence 12, Appl	Sequence 12, Appl	Sequence 12, Appl
4	464	32.4	334	4	US-09-454-279-21	Sequence 21, Appl	Sequence 21, Appl	Sequence 21, Appl	Sequence 21, Appl
5	423	29.5	261	4	US-09-454-279-16	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl
6	399	27.8	265	4	US-09-454-279-22	Sequence 22, Appl	Sequence 22, Appl	Sequence 22, Appl	Sequence 22, Appl
7	313	21.8	207	4	US-09-454-279-18	Sequence 18, Appl	Sequence 18, Appl	Sequence 18, Appl	Sequence 18, Appl
8	221	15.4	157	4	US-09-454-279-8	Sequence 8, Appl	Sequence 8, Appl	Sequence 8, Appl	Sequence 8, Appl
9	197.5	13.8	132	4	US-09-454-279-6	Sequence 6, Appl	Sequence 6, Appl	Sequence 6, Appl	Sequence 6, Appl
10	132	9.3	93	4	US-09-454-279-2	Sequence 2, Appl	Sequence 2, Appl	Sequence 2, Appl	Sequence 2, Appl
11	122	8.5	62	4	US-09-454-279-4	Sequence 4, Appl	Sequence 4, Appl	Sequence 4, Appl	Sequence 4, Appl
12	94	6.6	391	4	US-09-270-677-22309	Sequence 42309, A	Sequence 42309, A	Sequence 42309, A	Sequence 42309, A
13	94	6.6	502	4	US-09-635-872A-13	Sequence 13, Appl	Sequence 13, Appl	Sequence 13, Appl	Sequence 13, Appl
14	94	6.6	502	4	US-09-635-077A-13	Sequence 13, Appl	Sequence 13, Appl	Sequence 13, Appl	Sequence 13, Appl
15	94	6.6	502	4	US-09-636-060C-13	Sequence 13, Appl	Sequence 13, Appl	Sequence 13, Appl	Sequence 13, Appl
16	94	6.6	502	4	US-09-986-552-13	Sequence 13, Appl	Sequence 13, Appl	Sequence 13, Appl	Sequence 13, Appl
17	94	6.6	502	4	US-09-636-596C-13	Sequence 13, Appl	Sequence 13, Appl	Sequence 13, Appl	Sequence 13, Appl
18	94	6.6	502	4	US-10-023-894-16	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl
19	94	6.6	502	4	US-10-306-686-13	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl
20	92	6.4	717	4	US-09-583-110-4629	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl
21	92	6.4	721	4	US-09-107-433-2019	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl
22	90	6.3	292	4	US-09-107-532A-3741	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl
23	89.5	6.2	640	4	US-09-248-796A-16129	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl
24	89	6.2	335	4	US-09-934-901-6	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl
25	89	6.2	335	4	US-09-934-868-16	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl
26	89	6.2	335	4	US-10-321-2810-5	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl
27	89	6.2	335	4	US-10-320-8874-6	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl
Searched: 513545 seqs, 74659064 residues									
Total number of hits satisfying chosen parameters: 513545									
Minimum DB seq length: 0									
Maximum DB seq length: 2000000000									
Post-processing: Minimum Match 0%									
Listings first 45 summaries									
Database : Issued Patents AA:*									
1: /cgn2_6/ptodata/1/iaa/5A-COMB.pep:*									
2: /cgn2_6/ptodata/1/iaa/5B-COMB.pep:*									
3: /cgn2_6/ptodata/1/iaa/6A-COMB.pep:*									
4: /cgn2_6/ptodata/1/iaa/6B-COMB.pep:*									
5: /cgn2_6/ptodata/1/iaa/PCUTUS-COMB.pep:*									
6: /cgn2_6/ptodata/1/iaa/backfilesl1.pep:*									
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.									
SUMMARIES									
Result No. Score Query Length DB ID Description									
1	921	64.2	315	4	US-09-248-796A-18213	Sequence 18213, A	Sequence 18213, A	RESULT 1	US-09-248-796A-18213
2	508.5	35.5	45	4	US-09-248-796A-18213	Sequence 18213, A	Sequence 18213, A	Sequence 18213, A	Sequence 18213, A
3	481.5	33.6	331	4	US-09-454-279-12	Sequence 12, Appl	Sequence 12, Appl	Sequence 12, Appl	Sequence 12, Appl
4	464	32.4	334	4	US-09-454-279-21	Sequence 21, Appl	Sequence 21, Appl	Sequence 21, Appl	Sequence 21, Appl
5	423	29.5	261	4	US-09-454-279-16	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl	Sequence 16, Appl
6	399	27.8	265	4	US-09-454-279-22	Sequence 22, Appl	Sequence 22, Appl	Sequence 22, Appl	Sequence 22, Appl
7	313	21.8	207	4	US-09-454-279-18	Sequence 18, Appl	Sequence 18, Appl	Sequence 18, Appl	Sequence 18, Appl
8	221	15.4	157	4	US-09-454-279-8	Sequence 8, Appl	Sequence 8, Appl	Sequence 8, Appl	Sequence 8, Appl
9	197.5	13.8	132	4	US-09-454-279-6	Sequence 6, Appl	Sequence 6, Appl	Sequence 6, Appl	Sequence 6, Appl
10	132	9.3	93</						

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on:

March 31, 2005, 10:10:35 ; Search time 26 Seconds
(without alignments)
1036.181 Million cell updates/sec

Title: US-10-042-059B-2
Perfect score: 1434
Sequence: 1 MDFMKPETYVLDLGNIKDALV.....DDATQKGGYVDRFLSSGLY 280

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79;*

1: pir1;*
2: pir2;*
3: pir3;*
4: pir4;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	776.5	54.1	256	2	AM5921 chorimate mutase (EC 5.4.99.5) - yeast (Saccharomyces cerevisiae)
2	558.5	38.9	251	2	T37784 probable chorismate mutase
3	464	32.4	334	2	S28958 chorimate mutase
4	399	27.8	265	2	TR0796 chorimate mutase
5	166	11.6	82	2	T14902 chorimate mutase
6	119	8.3	70	2	T14901 chorimate mutase
7	114.5	8.0	469	2	CT0109 hypothetical prote
8	104	7.3	313	2	BP15138 tetrahydroperoxyl
9	104	7.3	336	2	SO0739 site-specific DNA
10	103.5	7.2	1732	2	T14039 protein kinase (EC
11	99	6.9	1033	2	T37715 actin-interacting
12	98	6.8	1702	2	T14050 protein kinase (EC
13	97.5	6.8	1286	2	TR7366 RNA polymerase sub
14	97	6.8	323	2	D90527 glycerol-3-phospho
15	97	6.8	1167	2	ED6963 DNA polymerase III
16	96.5	6.7	829	2	S2366 DNA topoisomerase
17	96.5	6.7	932	2	FB4465 hypothetical prote
18	96	6.7	808	2	HR4474 hypothetical prote
19	96	6.7	1286	1	HE6845 DNA-directed RNA p
20	96	6.7	1286	2	TR8521 DNA-directed RNA p
21	96	6.7	1286	2	AT7161 M6R protein - vari
22	95.5	6.7	560	1	FB9059 arginine-tRNA ligase
23	94.5	6.6	790	2	TR9683 hypothetical prote
24	94.5	6.6	1287	1	RN7247 DNA-directed RNA p
25	94	6.6	7829	2	TR5789 hypothetical prote
26	94	6.6	311	1	C69952 probable ribosomal
27	94	6.6	502	2	AF3110 probable ribosomal
28	94	6.6	612	2	SP3611 glucose inhibited
29	93	6.5	307	2	AV5020 hypothetical prote

ALIGNMENTS

Query	Match	Length	DB	Score	DB ID	Description
1	54.1%	256	2	AM5921	chorimate mutase (EC 5.4.99.5) - yeast (Saccharomyces cerevisiae)	
2	58.7%	251	2	T37784	probable chorismate mutase	
3	32.4%	334	2	S28958	chorismate mutase	
4	27.8%	265	2	TR0796	chorismate mutase	
5	11.6%	82	2	T14902	chorismate mutase	
6	8.3%	70	2	T14901	chorismate mutase	
7	114.5%	469	2	CT0109	hypothetical prote	
8	104%	313	2	BP15138	tetrahydroperoxyl	
9	104%	336	2	SO0739	site-specific DNA	
10	103.5%	1732	2	T14039	protein kinase (EC	
11	99%	1033	2	T37715	actin-interacting	
12	98%	1702	2	T14050	protein kinase (EC	
13	97.5%	1286	2	TR7366	RNA polymerase sub	
14	97%	323	2	D90527	glycerol-3-phospho	
15	97%	1167	2	ED6963	DNA polymerase III	
16	96.5%	829	2	S2366	DNA topoisomerase	
17	96.5%	932	2	FB4465	hypothetical prote	
18	96%	808	2	HR4474	hypothetical prote	
19	96%	1286	1	HE6845	DNA-directed RNA p	
20	96%	6.7	1286	2	TR8521	DNA-directed RNA p
21	96%	6.7	1286	2	AT7161	M6R protein - vari
22	95.5%	6.7	560	1	FB9059	arginine-tRNA ligase
23	94.5%	6.6	790	2	TR9683	hypothetical prote
24	94.5%	6.6	1287	1	RN7247	DNA-directed RNA p
25	94	6.6	7829	2	TR5789	hypothetical prote
26	94	6.6	311	1	C69952	probable ribosomal
27	94	6.6	502	2	AF3110	probable ribosomal
28	94	6.6	612	2	SP3611	glucose inhibited
29	93	6.5	307	2	AV5020	hypothetical prote

Scoring table: BLOSUM62									
Gapop 10.0 , Gapext: 0.5									
Searched: 1612378 seqs, 512079187 residues									
Total number of hits satisfying chosen parameters: 1612378									
Minimum DB seq length: 0									
Maximum DB seq length: 200000000									
Post-processing: Minimum Match 0%									
Maximum Match 100%									
Listing first 45 summaries									
Database : UniProt: 03:*									
1: uniprot_sprot:*									
2: uniprot_trembl:*									
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.									
Result No. Query Match length DB ID Description									
1	1434	100.0	:280	2	Q9P4DB	Q9p4db	pichia angu		
2	907	63.2	:267	2	Q6BK64	Q6bk64	debaromyce		
3	808	56.3	:256	2	Q6C5J7	Q6c5j7	yarroyo		
4	776.5	54.1	:256	1	CHMU YEAST	CHMU YEAST	saccharomy		
5	752.5	52.5	:260	2	Q6CVY3	Q6cvy3	kluyveromy		
6	719	50.1	:260	2	Q6FLZ7	Q6flz7	candida gla		
7	702.5	49.0	:259	2	Q7SBG5	Q7sbg5	ashbya goss		
8	594	41.4	:269	2	Q7S8R4	Q7s8r4	neurospora		
9	581.5	40.6	:263	2	Q6L8Q0	Q6l8q0	rosellinia		
10	579	40.4	:267	2	Q9YB2	Q9yb2	emericella		
11	558.5	38.9	:251	1	CHMU SCHPO	CHMU SCHPO	013739	schizophac	
12	508.5	35.5	:294	2	Q9YV60	Q9yv60	oryza sativ		
13	481.5	33.6	:316	2	Q9C544	Q9c544	arabidopsis		
14	475	33.1	:295	2	Q96VZ8	Q96vz8	cryptococcus		
15	466	32.5	:334	1	CHMU ARATH	CHMU ARATH	p42738	arabidopsis	
16	465	32.4	:255	2	Q9STB2	Q9stb2	lycopersico		
17	464.5	32.4	:316	2	Q9XP60	Q9xp60	arabidopsis		
18	408.5	28.5	:255	2	Q9HF89	Q9hf89	oryza sativ		
19	399	27.8	:265	2	Q9S7H4	Q9s7h4	arabidopsis		
20	391	27.3	:284	2	Q6Z9B6	Q6z9b6	oryza sativ		
21	290.5	20.3	:179	2	Q6FP31	Q6fp31	nicotiana		
22	247	17.2	:131	2	Q65B61	Q65b61	prunus aviu		
23	166	11.6	:82	2	Q22410	Q22410	petroselinu		
24	119	8.3	:70	2	Q22409	Q22409	petroselinu		
25	114.5	8.0	:469	2	Q51102	Q51102	borrelia bu		
26	111	7.7	:637	2	Q651H3	Q651h3	bacillus li		
27	107	7.5	:622	2	Q7NAK6	Q7nak6	mycoplasma		
28	106.5	7.4	:413	2	Q8R3P7	Q8r3p7	mus musculu		
29	106.5	7.4	:1286	2	Q8V424	Q8v424	monkeypox		
30	106	7.4	:5289	2	Q88033	Q88033	lactobacill		
31	105.5	7.4	:1638	2	Q8IWQ7	Q8iwq7	homo sapien		

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model.
Run on: March 31, 2005, 10:21:29 ; Search time 4655 Seconds
(without alignments)
2314.599 Million cell updates/sec

Title: US-10-042-059B-2
Perfect score: 1434

Sequence: 1 MDFMKPSTVLDLGNIRDALV.....DDATQKSGGYVDRFLSSGLY 280

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues.

Total number of hits satisfying chosen parameters: 9416466
Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n: model -DEV=x1h
-Q/cgn2_1/USP0/spool/US010042059/runat_30032005_103454_9002/app/query.fasta_1.455
-DB=GenEmbl -ONMT=fastap -SUFFIX=reg -MINMATCH=0.1 -LOCPLIC=0 -LOCPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=bleum62 -TRANS=human40.cdt -LIST=45
-DOCALIGN=200 -THR SCOREPECT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-ONMT=PO -NORMnext -HMAPSIZE=500 -MINLEN=000000000
-USER=US010042059 @CEN 1 1 400 @runat_J0032005_103454_9002 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -LONGLOG
-DEV TIMEOUT=120 -WARN TMBOND=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:
1: gb_bac:
2: gb_hhg:
3: gb_in:
4: gb_on:
5: gb_ov:
6: gb_pat:
7: gb_phi:
8: gb_pil:
9: gb_pr:
10: gb_ro:
11: gb_sts:
12: gb_sy:
13: gb_uu:
14: gb_vl:
Pred. No. 18 is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is greater by analysis of the total score distribution.

ALIGNMENTS

RESULT 1
BD266599
LOCUS
DEFINITION Nucleic acid molecule, containing a nucleic acid which codes for a polyPeptide with chorismate mutase activity.
ACCESSION BD266599
VERSION BD266599.1
KEYWORDS
SOURCE
ORGANISM Pichia angusta
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Pichia.
REFERENCE 1 (bases 1 to 833)
AUTHORS Geilissen, G., Braus, G., Pries, R., Krappmann, S. and Strasser, A. W.
TITLE Nucleic acid molecule, containing a nucleic acid which codes for a polyPeptide with chorismate mutase activity
PATENT: JP 20052788-A 1 17-DEC-2002;
RHEIN BIOTECH GESELLSCHAFT FUER NEUE BIOTECHNOLOGISCHE PROZESSE UND
PRODUKTE MBH
COMMENT OS Hansenula polymorpha (yeast)
PN JP 20052788-A/1
PD 17-DEC-2002
PP 27-APR-2000 JP 200614405
PR 27-APR-1999 DE 199 19 124-7
PI GERD GEILISSEN, GERHARD BRAUS, RALPH PRIES, SVEN KRAPPmann, PI
ALEXANDER W STRASSER
PC C12N15/09, C07K16/40, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC
C12N9/90, C12N15/00,

RESULTS

No.	Score	Query	Match Length	DB	ID	Description
1	1434	100.0	843	6	BD266599	BD266599 Nucleic a
2	1434	100.0	843	6	AX043758	AX043758 Sequence
3	1434	100.0	1652	8	AF204738	AF204738 Pichia an
4	1434	100.0	1655	6	BD266600	BD266600 Nucleic a

SUMMARIES

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame plus p2n model
Run on: March 31, 2005, 10:21:01 ; Search time 584 Seconds
(without alignments)
2838.232 Million cell updates/sec

Title: US-10-042-059B-2

Perfect score: 1434
Sequence: 1 MDFWKRPETVLDLGNIRDALV.....DDATQKSGGYVDRFLSGLY 280

Scoring table: BLOSUM62

Xgapop 1.0.0 , Xgapext 0.5
Ygapop 1.0.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delect 7.0

Searched: 439026 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412
Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODS=frame+P2N:modelf -DBV=x1h
-Q=/sgn2_1/USPRO_spool/US1042059/runat_30032005_103453_8996/app_query.fasta_1.455
-DBN=Geneseq_16Dec04 -QFM=factcap -SURFIX=rng -MINMATCH=0.1 -LOCPCL=0
-LOOPEXT=0 -UNITS=b16 -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=PCP -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODS=LOCAL -OUTFMT=txt -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US1042059 @CGN_1_644 @runat_30032005_103453_8996 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQTY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=1200 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DDELOP=6 -DDELEXT=7

Database :

N_Geneseq_16Dec04: *
1: geneseq1908: *
2: geneseq1930s: *
3: geneseq2000s: *
4: geneseq2001as: *
5: geneseq2001bs: *
6: geneseq2002as: *
7: geneseq2002bs: *
8: geneseq2003as: *
9: geneseq2003bs: *
10: geneseq2003bs: *
11: geneseq2003ds: *
12: geneseq2004as: *
13: geneseq2004bs: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	DB ID	Description
1	1434	100.0	843	3	AAC81949
2	1434	100.0	155	3	AAC81950
3	921	64.2	807	12	ADP98547
4	587.5	41.0	801	13	ADR85510
5	540.5	37.7	869	8	AAL150202

ALIGNMENTS

RESULT 1					
AAC81949	ID	AAC81949	standard; DNA; 843 BP.	XX	
XX				AC	AAC81949;
XX				DT	15-SBR-2003 (revised)
XX				DT	01-MAR-2001 (first entry)
XX				DE	H. polymorpha chorismate mutase DNA.
XX				DE	KW Chorismate mutase; prephenate; selection marker; auxotrophic yeast; ds.
XX				OS	Pichia angusta.
XX				PN	WO20065071-A1.
XX				XX	
XX				PD	02-NOV-2000.
XX				PR	27-APR-2000; 2000WO-EP003844.
XX				PR	27-APR-1999; 99DB-01019124.
XX				PA	(RHEI-) RHEIN BIOTECH GES NEUB BIOTECHNOLOGISCHE.
XX				PI	Gellissen G, Braus G, Pries R, Krappmann S, Strasser AW;
XX				DR	WPI: 2000-687355/67.
XX				DR	P-PSDB; AAB11451.
XX					AAB150202 M sterilis

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Om protein - nucleic search, using frame_plus_p2n model

Run on: March 31, 2005, 10:42:01 ; Search time 195 Seconds

(without alignments)
2349.526 Million cell updates/sec

Title: US-10-042-059B-2

Perfect score: 1434

Sequence: 1 MDFMKPETYLDLGNIRDALV..... DDAQKSGGYVDRPLISGLY 280

Scoring table: BLOSUM62

Xgapop 1.0 , Xgapext 0.5

Ygapop 1.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODLU=frame+p2n:model -DRV=x1h
-Q=/cgn2_1/USPTO_spool/US10042059/runat_30032005_103456_9019/app_query.fasta_1.455
-O=/cgn2_1/USPTO_spool/US10042059/runat_30032005_103456_9019/app_query.fasta_1.455
-DB=issued_patents_NA -QFM=fastap -SUFFIX=rni -MINMATCH=0.1 -LOCPL=0
-LOCPXT=0 -UNITS=5ts -START=1 -END=-1 -MATRIX=blosum62 -TRANS=unamend0.cdl
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODLU=LOCAL -OUTFMT=PTO -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10042059 @CGN 1 1.177 @runat_30032005_103456_9019 -NCPU=5 -ICFU=3
-NO MAP -LARGEOUT -NEGSORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A/COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5B/COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A/COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B/COMB.seq:*

5: /cgn2_6/ptodata/1/ina/PCUTS/COMB.seq:*

6: /cgn2_6/prodata/1/ina/backfile1.seq:*

Pred: No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

1	921	64.2	948	4	US-09-248-796A-4110	Sequence 4110, Appl	8.57e-116	Length: 948	Sequence 3, Appl
2	508.5	35.5	1231	4	US-09-454-229-13	Sequence 13, Appl	921.00	Matches: 170	Sequence 9, Appl
3	481.5	33.6	1223	4	US-09-454-229-11	Sequence 11, Appl	82.63	Conservative: 44	Sequence 9, Appl
4	464.5	32.4	1217	3	US-09-610-040-4	Sequence 4, Appl	65.64	Mismatches: 45	Sequence 17, Appl
5	464.5	32.4	1217	3	US-09-610-040-10	Sequence 10, Appl	64.23	Index: 0	Sequence 5, Appl
6	464.5	32.4	1217	4	US-10-267-763-4	Sequence 4, Appl	Gaps: 0		Sequence 6, Appl
7	464.5	32.4	1217	4	US-10-267-763-10	Sequence 10, Appl			Sequence 3, Appl
8	464	32.4	1207	3	US-09-610-040-2	Sequence 2, Appl			Sequence 9, Appl
9	464	32.4	1207	3	US-09-610-040-8	Sequence 8, Appl			Sequence 7, Appl
10	464	32.4	1207	4	US-10-267-763-2	Sequence 2, Appl			Sequence 5, Appl
11	464	32.4	1207	4	US-10-267-763-8	Sequence 8, Appl			Sequence 6, Appl
12	423	29.5	1020	4	US-09-454-229-15	Sequence 15, Appl			Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-09-248-796A-4110
; Sequence 4110, Application US/09248-796A

; Patent No. 6747137

; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al

; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBIC

; FILE REFERENCE: 10196.132

; CURRENT APPLICATION NUMBER: US/09/248-796A

; CURRENT FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; PRIOR FILING DATE: 1998-06-13

; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO: 4110

; LENGTH: 948

; TYPE: DNA

; ORGANISM: Candida albicans

; US-09-248-796A-4110

; Alignment Scores:

Pred. No.: 8.57e-116
Score: 921.00
Percent Similarity: 82.63%
Best Local Similarity: 65.64%
Query Match: 64.23%

DB: 4

Gaps: 0

US-10-042-059B-2 (1-280) x US-09-248-796A-4110 (1-948)

Qry: 1 MetapphenetyllysProglutylvalleuspleuArgValleuVal 20
DB: 142 ATGGATTATGAAACCGAACTGTGTGATCTGCCRACACCCGTCAGGTTGGTA 201

GenCore version 5.1.6
(c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: March 31, 2005, 11:54:44 ; Search time 622 Seconds
2724.594 Million cell updates/sec

Title: US-10-042-059B-2
Perfect score: 1434

Sequence: 1 MDFMKETVLDLGNRIDLALV.....DDATOKSGG3YVDRFLSGLY 280

Scoring table: BLASTM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delex 7.0

Searched: 5607317 seqs, 3026245999 residues

Total number of hits satisfying chosen parameters: 11214634

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODBL=frame+pn2n,model -DEV=x1h
-Q=/cgn2_1/USPRO.spool/US10042059/runat_30032005_103458_9071/app/query.fasta_1-455

-DB=Published Applications -QFMT=app -MINMATCH=0.1
-LOOPCL=0 -LOOPENEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blastm62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAXK=100

-THR MINN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZES=500 -MININTN=0
-MAXINTN=2000000000 -USER=US10042059 @CGN_1_1_688 @runat_30032005_103458_9071

-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=0 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEX=7

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pupnna/us07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/1/pupnna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/1/pupnna/us06_PUBCOMB.seq:*

4: /cgn2_6/ptodata/1/pupnna/us07_NEW_PUB.seq:*

5: /cgn2_6/ptodata/1/pupnna/PCUTS_PUBCOMB.seq:*

6: /cgn2_6/ptodata/1/pupnna/us08_NEW_PUB.seq:*

7: /cgn2_6/ptodata/1/pupnna/us09_NEW_PUB.seq:*

8: /cgn2_6/ptodata/1/pupnna/us09_PUBCOMB.seq:*

9: /cgn2_6/ptodata/1/pupnna/us09_PUBCOMB.seq:*

10: /cgn2_6/ptodata/1/pupnna/us09C_PUBCOMB.seq:*

11: /cgn2_6/ptodata/1/pupnna/us09_PUBCOMB.seq:*

12: /cgn2_6/ptodata/1/pupnna/us09c_NEW_PUB.seq:*

13: /cgn2_6/ptodata/1/pupnna/us10_PUBCOMB.seq:*

14: /cgn2_6/ptodata/1/pupnna/us10_PUBCOMB.seq:*

15: /cgn2_6/ptodata/1/pupnna/us10c_PUBCOMB.seq:*

16: /cgn2_6/ptodata/1/pupnna/us10c_PUBCOMB.seq:*

17: /cgn2_6/ptodata/1/pupnna/us10c_PUBCOMB.seq:*

18: /cgn2_6/ptodata/1/pupnna/us10c_PUBCOMB.seq:*

19: /cgn2_6/ptodata/1/pupnna/us09c_PUBCOMB.seq:*

20: /cgn2_6/ptodata/1/pupnna/us11c_PUBCOMB.seq:*

21: /cgn2_6/ptodata/1/pupnna/us11c_PUBCOMB.seq:*

22: /cgn2_6/ptodata/1/pupnna/us60_PUBCOMB.seq:*

Result No.	Score	Query Match Length	DB ID	Description
1	1434	100.0	843	Sequence 1, Appli
2	1434	100.0	1655	Sequence 3, Appli
3	921	64.2	807	Sequence 6032, Appli
4	540.5	37.7	869	Sequence 26, Appli
5	517.5	36.1	2920	Sequence 249, Appli
6	517.5	36.1	2987	Sequence 5249, Appli
7	508.5	35.5	1016	Sequence 93956, Appli
8	508.5	35.5	1231	Sequence 13, Appli
9	508.5	35.5	1231	Sequence 13, Appli
10	500.5	34.9	1548	Sequence 4455, Appli
11	497.5	34.7	1146	Sequence 7285, Appli
12	495.5	34.5	1423	Sequence 123820, Appli
13	489.5	34.1	1376	Sequence 102347, Appli
14	486	33.9	1500	Sequence 14353, Appli
15	481.5	33.6	1223	Sequence 11, Appli
16	481.5	33.6	1223	Sequence 11, Appli
17	479.5	33.4	1430	Sequence 737, Appli
18	478.5	33.4	1397	Sequence 4454, Appli
19	470.5	32.8	658	Sequence 102346, Appli
20	469.5	32.7	1368	Sequence 6992, Appli
21	466	32.5	1005	Sequence 102349, Appli
22	466	32.5	1005	Sequence 737, Appli
23	464.5	32.4	1217	Sequence 10, Appli
24	464	32.4	1217	Sequence 2, Appli
25	464	32.4	1207	Sequence 8, Appli
26	464	32.4	1207	Sequence 15, Appli
27	451	31.8	1299	Sequence 15, Appli
28	423	29.5	1020	Sequence 15, Appli
29	423	29.5	1020	Sequence 15, Appli
30	423	29.5	1020	Sequence 15, Appli
31	419	29.2	1653	Sequence 93056, Appli
32	408.5	28.5	768	Sequence 49372, Appli
33	399	27.8	798	Sequence 972, Appli
34	399	27.8	798	Sequence 972, Appli
35	399	27.8	1006	Sequence 3, Appli
36	399	27.8	1006	Sequence 9, Appli
37	385	26.8	1010	Sequence 102351, Appli
38	381	26.6	1314	Sequence 15170, Appli
39	374.5	26.1	1066	Sequence 1332, Appli
40	374.5	26.1	1083	Sequence 15559, Appli
41	373	26.0	1197	Sequence 27576, Appli
42	373	26.0	921	Sequence 83364, Appli
43	355	24.8	622	Sequence 4913, Appli
44	351.5	24.5	2154	Sequence 83551, Appli
45	346	24.1	580	Sequence 12291, Appli
ALIGNMENTS				

RESULT 1

US-10-042-059A-1

Sequence 1, Application US/10042059A

Publication No. US2002019704A1

GENERAL INFORMATION:

APPLICANT: Rhein Biotech Gesellschaft fur neue biotechnologische Prozesse und

ANPPLICANT: Produkte mBH

TITLE OF INVENTION: Nucleic Acid Molecule Containing a Nucleic Acid Coding for a Pol

TITLE OF INVENTION: with Chaperonate Mutase Activity

FILE REFERENCE: P30589US019

CURRENT APPLICATION NUMBER: US/10/042, 059A

CURRENT FILING DATE: 2001-10-25

PRIOR APPLICATION NUMBER: DE 199 19 124 7

PRIOR FILING DATE: 1999-04-27

NUMBER OF SEQ ID NOS: 3

SEQUENCE LENGTH: 843

ORGANISM: Hansenula polymorpha

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

GenCore version 5.1.6
(c) 1993 - 2005 Compugen Ltd.

On protein - nucleic search, using frame_plus_p2n model
Run on: March 31, 2005, 10:22:54 ; Search time 7476 Seconds
(without alignments)
1425.628 Million cell updates/sec

Title: US-10-042-059B-2
Perfect score: 1434

Sequence: 1 MDPFMKPETVLDLGNIRDALV.....DDATQKGGYVDRPLSSGLY 280

Scoring table: BLOSUM62
Xgapext 10.0 , Ygapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delett 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088
Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MDEBLI=frame+P2n;model -DEV=x1h
-Q-/cgn2.1/USP0_spool/US10042059/runat_30032005_103455_9011/app_query.fasta_1.455
-DDE=EST -QFMT=FASTA -SURPCPT=EST -THR MAX=100 -THR MIN=0 -ALIGNM=0 -ALIGNM=15 -NODEL=LOCAL
-OUTPNT=EST -NORM=ext -HRSPIRE=500 -MINLEN=2000000000
-USER=US10042059@cgn.1.1.352@runat_30032005_103455_9011 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: gb_east1:
2: gb_east2:
3: gb_east3:
4: gb_east4:
5: gb_east5:
7: gb_east6:
8: gb_gb81:
9: gb_gb82:
Yarrowia lipolytica

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	758	52.9	1185	CNS06TSP
2	570	39.7	7	CO005472
3	506.5	35.3	574	BX303243
4	481.5	33.6	1222	AY103806
5	472	32.9	7	CO165801
6	470	32.8	1329	BX822123
7	469	32.7	937	CNS0A727
8	469	32.7	10533	AL41049
9	465	32.4	1265	CNS0A6W

RESULT	1	CNS06TSP
LOCUS	CNS06TSP	1185 bp DNA linear
DEFINITION	T7 end of clone AWOA029A11 of library AWOA from strain CLIB 89 of Yarrowia lipolytica, genomic survey sequence.	GSS 05-011-2001
ACCESSION	AL414131	
VERSION	AL414131.1	GI:12186949
KEYWORDS	Yarrowia	
SOURCE	Yarrowia lipolytica	
ORGANISM	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Diopodascaceae; Yarrowia.	
REFERENCE	1 (bases 1 to 1185)	
AUTHORS	Souciet,J.L., Aigle,M., Artiguenave,P., Blandin,G.J., Boletin-Fukuhara,M., Bon,E., Bröttier,P., Casarego,J.S., de-Montigny,J., Dujon,B., Durran,P., Lepingle,A., Llorente,B., Malperruy,A., Neuveglise,C., Orléa,Kalogeropoulos,O., Potier,S., Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M., Wincker,P. and Weissenbach,J.	
TITLE	Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies	
JOURNAL	FEBS Lett. 487 (1), 3-12 (2000)	
MEDLINE	20584711	
PubMed	11152876	
REFERENCE	2 (bases 1 to 1185)	
AUTHORS	Casarego,S., Neuveglise,C., Lepingle,A., Bon,E., Feyerol,C., Artiguenave,F., Wincker,P. and Gailardin,C.	
TITLE	Genomic exploration of the hemiascomycetous yeasts: 17. Yarrowia lipolytica	
JOURNAL	FEBS Lett. 487 (1), 95-100 (2000)	
MEDLINE	20584727	
PubMed	11152892	
REFERENCE	3 (bases 1 to 1185)	